

## **Explanatory Integration**

**Andrew Wayne**

**Department of Philosophy, University of Guelph**

**Forthcoming in *European Journal for Philosophy of Science***

### **Abstract**

The goal of this paper is to show how scientific explanation functions in the context of idealized models. It argues that the aspect of explanation most urgently requiring investigation is the nature of the connection between global theories and explanatory local models. This aspect is neglected in traditional accounts of explanation (Hempel 1965; Kitcher 1989). The paper examines causal, minimal model, and structural accounts of model-based explanation (Woodward 2003; Batterman and Rice 2014; Bokulich 2011). It argues that they too fail to offer an account of the connection with global theory that can justify the explanatory power of an idealized local model, and consequently these accounts are unable effectively to distinguish explanatory from non-explanatory models. On the account proposed here, scientific explanation requires theoretical integration between the local model described in the explanation and a global theory with independent explanatory power.

Keywords: Explanation; model; science; idealization

Word count: 10,138 words

## 1. Introduction

A decisive moment in the analysis of scientific explanation occurred in 1948, when Carl Hempel and Paul Oppenheim published the first thoroughly empiricist account of explanation in science (Hempel and Oppenheim 1948). Their deductive-nomological account of explanation was developed by Hempel in the 1960s (Hempel 1965) and expanded under the rubric of “explanatory unification” by Philip Kitcher in the 1980s (Kitcher 1981; Kitcher 1989). Today, this non-causal, deductivist approach to scientific explanation has fallen on hard times. At worst, deductivism faces a host of counterexamples. At best, it seems simply irrelevant to the explanatory practices of contemporary science.

The deductivist approach merits another look. It will be particularly helpful, I suggest, in the context of explanations appealing to highly idealized models. Idealized models contain elements that do not accurately represent components of the target system itself. Recent work in philosophy of science has argued that these elements of idealized models may nonetheless be used to explain phenomena or regularities in the target (Woodward 2003; Strevens 2008; Batterman 2010; Bokulich 2011; Batterman and Rice 2014; Healey 2015; Reutlinger 2016; Saatsi 2016; [self-reference omitted]). The goal of this paper is to show how scientific explanation functions in the context of idealized models by articulating certain local and global necessary conditions that the explanation and the models must satisfy. I call this account of scientific explanation *explanatory integration*. Section 2 distinguishes local and global features of explanation, and briefly describes the local features. Section 3 explores the global features and examines Kitcher’s criterion that explanation requires unification. Kitcher’s account faces serious problems, but this does not show that the deductivist approach is unworkable. Rather, we need to change the focus slightly. The key challenge is not that of developing formal conditions on the degree of unification of global theories, as Kitcher tried to do. Rather, the aspect of explanation most urgently requiring investigation is one neglected by earlier deductivists: the nature of the connection between global theories and explanatory local models. Section 4 reviews prominent causal, minimal model, and structural accounts of model-based explanation (Woodward 2003; Batterman and Rice 2014; Bokulich 2011). It argues that they too fail to offer an account of the connection with global theory that can justify the explanatory power of an idealized local model, and consequently these accounts are unable effectively to distinguish explanatory from non-explanatory models. Section 5 presents the outlines of such an account. The central idea is that scientific explanation requires theoretical integration between the local model described in the explanation and a global theory with independent explanatory power.

## 2. Local features

An account of scientific explanation may be characterized by two types of requirements any good explanation must meet. The first concerns local features of an explanation, what Hempel and Oppenheim call “conditions of adequacy” (Hempel and Oppenheim 1948, 247). These

conditions have to do with characteristics of the explanans and the relation between the explanans and the explanandum. The second type concerns global features of an explanation, the relation between the explanation and the larger scientific field in which it is embedded. Sections 3 and 5 focus on global conditions. This section explores local conditions.

On a deductivist approach, the basic local condition is that the explanandum be entailed by the explanans. This ensures that the occurrence of the phenomenon or regularity described in the explanandum was to be expected, given the empirical conditions and relevant laws or dependency relations stated in the explanans. In the deductive-nomological (D-N) account, local conditions also include, roughly speaking, that the explanans contain one or more laws, that these laws be essential for the explanation, that the explanans have empirical content, and that the premises in the explanans be true or approximately true (Hempel and Oppenheim 1948, 247-249). A major challenge for deductivist approaches to explanation is to offer a precise articulation of these local conditions (Kim 1963; Hempel 1965; Kitcher 1989). Common wisdom has it that deductivists failed decisively in this project, especially in the context of model-based explanation.

The challenge here is that Hempel and Kitcher take the statements in the explanans and explanandum to be true or approximately true of the system that is the target of the explanation. For explanations making use of idealized models, however, this is not the case. In these sorts of explanations, the premises in the explanans are true of a model, but they may not correspond to features of the target system, even approximately. The local model, then, is the truthmaker for the premises in an explanatory derivation. In cases where the model has elements that do not accurately represent objects or phenomena in the target system—the cases that are our focus—some premises in the explanans are not even approximately true of the target system. The solution here is to require that premises in the explanans be true of a local model, not a target system. The explanandum statement, by contrast, must be true of the model *and* approximately true of the target system—since normally an explanation aims at a feature of the target system, not just a feature of the model.

Requiring that the explanans be true of the model, not the physical system, may seem like a false solution to the deductivists' problem. Model creation is guided by the interests of model builders, and so one might worry that models can be created with an arbitrary set of dependency relations. On the approach to explanation proposed here, explanation appeals to dependencies in a model, not dependencies in a target system. Why isn't it the case that anything can be made to explain anything? The answer lies in considering global features of explanation (Section 3). As we shall see, models underwriting explanations must be theoretically integrated with a global theory. This puts a constraint on explanatory models but leaves considerable latitude in model building to accommodate the interests of the model builders. Typical Newtonian models include dependency relations of the sort described in the next paragraph. For certain purposes, however, a model may be constructed in which mass does depend on applied force, and—as long as it

meets local and global conditions—such a model may be explanatory. Pragmatic considerations drive the construction of models: which target systems are of interest, which features of the system are to be modeled, and whether the model aims at predictive accuracy, simplicity, consilience with other models, etc. On the present account, however, once a model is built it is an objective fact about the model whether or not it can underwrite a given scientific explanation.

Another frequent criticism levelled at Hempel and Kitcher is that the deductivist requirement that the explanans subsume the explanandum under a law cannot be enough. This is the root of the most aggravating traditional counterexamples to deductivist approaches to explanation, the problems of explanatory asymmetry and irrelevance (recall flagpoles and samples of hexed salt). Closer to scientific practice, the law of conservation of energy, for example, enables one to deduce the trajectory of a Newtonian mass (given appropriate empirical conditions) in a Newtonian model. But it does not explain *why* the mass follows a given trajectory, say the critics. Rather, what is needed in the explanation is to show what the trajectory depends upon in the model. In a typical Newtonian model, it is true that if the applied force had been different, the trajectory of the point particle would have been different, but it is false that if the applied force had been different, the mass would have been different. Truthmakers for these claims are features of the model itself, and in particular its dependency relations. Thus, explanations need to include how the explanandum is governed by dependency relations in the model. To do this, the explanans must describe at least one relevant dependency relation in the model. This is a dependency relation among features of the model; sometimes, as we have seen, these are features that have no accurate real-world correlate.

The account of explanation developed here views models as abstract objects. This approach to models goes back to Ron Giere's three-part analysis of a scientific theory as a model, which is an abstract object; a set of propositions (axioms) that are true of the model; and a set of propositions that describe how the model is similar to or represents elements of the target system (Giere 1988). While Giere's general approach is presupposed in virtually all work on model-based explanation, there is a range of views about how to understand these abstract entities, and specifically whether they can be usefully understood as fictions, in analogy with literary fictions. Alisa Bokulich has a strong commitment to models as fictions, fictions that have significant explanatory power (Bokulich 2009; we return to Bokulich's account in Section 4). Giere, by contrast, argues that these abstract entities should be analyzed in terms of their similarity relations with physical systems, and relating them to fictions is a mistake (Giere 2009). The present account is compatible with the current range of views about the extent to which models are fictions.

The local conditions put forward here—that the statements be true of a model and that the explanans include dependency relations—raise a host of further questions. The issues of truthmakers for explanatory statements about models and the role of dependency relations in model-based explanation are challenging. Since models are abstract objects, modal claims about

models seem problematic in ways above and beyond the normal difficulties we have making sense of these sorts of claims about the physical world. However, these problems are not unique to my account, and all causal and non-causal approaches to model-based explanation face the same worries about making sense of modalities in idealized models. These problems lie outside the scope of the present paper. All I hope to have shown here is that to make progress on these issues, previously neglected global features of explanation must be better understood. To these we now turn.

### 3. Global features

What I am calling the global features of an explanation are features external to the explanation itself. These are not characteristics of the explanans alone, the explanandum, the relation between the two, or their relation to the target system. Rather, they concern the relation between the explanation—or more precisely, the model described by the explanation—and the larger scientific field in which it is embedded. The aim of this section is to explore what these global features of explanation are and how they have been treated in prior work on scientific explanation. As we shall see below and in Section 4, global features have been largely ignored in philosophical accounts of explanation, with the notable exception of work by Hempel and Kitcher.

First, however, a reasonable doubt must be addressed: Why do scientific explanations need satisfy global conditions at all? Prominent accounts of model-based explanation, such as those of James Woodward, Robert Batterman and Collin Rice (Woodward 2003; Batterman 2002; Batterman and Rice 2014), are couched entirely in terms of local features. As I shall argue, neglecting global features creates problems for these accounts (Section 4). The D-N account itself, the foundation of the deductivist tradition, is commonly taken not to include any global conditions on explanation. This, I believe, illustrates a common misunderstanding about the D-N account and about the role of global features in explanation more generally.

Admittedly, the core of the D-N account, its logical and empirical “general conditions of adequacy” for explanation, are entirely local conditions. And standard applications make no direct reference to global conditions: from the law that light travels in a straight line, the height of the flagpole and the angle of the sun, the length of the shadow can be predicted and thus explained. Yet this is not the whole story about the D-N account. The rest of the story, what Kitcher dubs the “unofficial” view of explanation, was never part of the formal D-N conditions but emerges clearly in the work of Hempel and others. As Hempel develops this idea, the explanatory power of a law statement in an explanans is underwritten by its connection to a broader and deeper scientific theory, where the latter theory “provides a systematically unified account of many different empirical laws” (Hempel 1965, 444). Call a broad and deep scientific theory a *global theory*, as for example the standard model in particle physics, the general theory of relativity, or contemporary evolutionary theory. Hempel is claiming that explanation has an

ineliminably global component. Whether a derivation is explanatory depends not only on local conditions but also on the connection between statements in the explanans, specifically laws, and a global scientific theory.

As Kitcher frames it, the unofficial view is that explanation requires unification, both in the sense that the explanation is unified with a global theory, and in the sense that the global theory is itself unified (Kitcher 1981). These assumptions are front and center in Kitcher's account, where the explanatory power of a derivation is determined by a set of formal criteria defining the degree of unification of the global theory in which the derivation is embedded. So we need to distinguish two kinds of global conditions on explanation: the right sort of connection between the global theory and local model described in the explanation, and the independent explanatory power of the global theory.

The former condition requires there be a connection between global theory and local model such that the global theory confers explanatory power on the local model. Hempel took this connection to be entirely straightforward: an empirical law in an explanans is simply identical with the global theory or deducible from the global theory plus empirical conditions described in the explanans. For example, a derivation of the period of a pendulum using Galileo's pendulum law is explanatory, according to Hempel, because this law is a deductive consequence of Newton's laws plus empirical assumptions, such as an infinitely small oscillation amplitude and constant gravitational force. Kitcher also took this connection to be simple and transparent. Like Hempel, Kitcher assumes that law statements featuring in explanations will be unproblematically derivable from—or identical with—a global theory. At worst, Kitcher allows that laws featuring in explanations may be straightforward and well-justified approximations of an overarching theory (Kitcher 1989, 452-454).

However, it is widely recognized that in the explanatory practice of contemporary science this connection need not be deductive or approximately deductive, and it is anything but simple. For many sorts of model-based explanations in science, the premises in the explanans describe features of a highly idealized local model. Rather than deduction, an intricate set of steps involving such procedures as abstractions, singular perturbation methods, and contrary-to-fact assumptions is required to get from the global theory to the law statement about the local model contained in the explanans (Batterman 2002; [self-reference omitted]). In a given field of science, there are general pathways of constructing these bridges from global theory to local model. But these pathways are not contained within the formalism of the global theory itself; the global theory does not provide an algorithm for constructing explanatory local models. Presumably, some of these pathways count as "appropriate connections" between the local model and the global theory such that the derivation describing the local model is an explanation. Other pathways do not. What is needed is a detailed analysis of the conditions under which these pathways support explanation. My contention is that this project has been neglected by Hempel and Kitcher—and, as we shall see, neglected in recent work on model-based explanation—yet it

is precisely the aspect of explanation most urgently requiring investigation. Indeed, the main goal of the present paper is precisely to explore and define these conditions (Section 5).

Let us return to the other global condition on explanation, that the global theory itself must have independent explanatory power. Implicit in Hempel's approach to explanation are the ideas that certain successful and well-confirmed global theories in mature sciences provide unification of some sort, that this unification is conducive to understanding and explanation, and that these global theories have explanatory power. With respect to what the features are of the global theory that give it explanatory power, however, Hempel remains (uncharacteristically for Hempel!) quite vague. He makes scattered remarks to the effect that global theories provided unification and unification is conducive to understanding. He does not venture to investigate further what precisely the features of global theories are that make them conducive to underwriting scientific explanation. Rather, he seems to take it as a matter of scientific consensus that, for example, Einstein's relativity theory or Darwinian evolutionary theory have explanatory power.

By contrast, the major successor to the D-N model, Kitcher's explanatory unification, is mainly focused on rendering precise the intuition that a global theory has independent explanatory power to the extent that it is unified. Kitcher's focus is on the intuitive idea that explanation and understanding have to do with science's ability to unify our account of empirical phenomena. Kitcher insists that a detailed set of conditions on the explanatory power of global theories can be given. His ambitious project is to cash out the explanatory power of global theories by defining degree of unification in terms of a set of general and formal criteria applicable to any global theory in science. Degree of unification is then the measure of the explanatory power of a mature and well-confirmed global theory. In this way, Kitcher's account of explanatory unification aims not just to articulate a key necessary condition on explanation but also to give deductivism the tools to respond to the challenges of explanatory irrelevance and explanatory asymmetry, among others, that bedevilled the D-N account.

Explanatory unification, however, faces significant problems of its own. Central to this account is the insight that one important component of a scientific theory is a set of argument patterns employed by scientists proficient in that theory. Distinct arguments offered by scientists in the course of making predictions or giving explanations share an argument pattern if they are similar in both logical structure and in the nomological vocabulary they use. The degree of unification of a theory is determined by examining these argument patterns. Roughly speaking, a unified theory is one that uses few argument patterns to explain a large range of results. The key difficulty of Kitcher's approach is how to define what counts as an argument pattern, where this definition must apply to a diverse range of global theories. According to Kitcher, the definition must enable us to compare the argument patterns, and hence degree of unification, of Darwinian evolutionary biology and Newtonian mechanics (Kitcher 1981). And it must be able to rule out cases like that of the spurious "unified" theory of biology, which derives all the results of

evolutionary biology from the single argument pattern "all living organisms are the product of descent with modification, therefore X" (where X is any statement constituting the conclusion of an explanation or prediction of evolutionary biology). Kitcher considers at length the problem of spurious unification, that is, defining argument patterns such that derivations are lumped together which our scientific common sense tells us are dissimilar. Kitcher's strategy is to develop quite complicated syntactic requirements on acceptable argument patterns in order to debar ever-more-clever gerrymandering (1981, 526-29; 1989, 480-88). Kitcher rescues his philosophical account of unification from (as he sees it) degenerating into toothlessness by legislating the syntactic form an acceptable argument pattern must have ([self-reference omitted]).

This strategy seems misguided, for we have good reason to be sceptical that any general syntactic rules governing acceptable argument patterns are to be found. All the derivations associated with a given theory will be similar to each other in a myriad of ways, but most of these similarities will be terribly uninteresting. (For example: the derivations make use of rational numbers, they invoke mathematical operators, they have steps which proceed by modus ponens, and so on). The crucial element of the concept of an argument pattern is that it picks out a *relevant* similarity among derivations. For this reason, Kitcher's strategy is unlikely to succeed ([self-reference omitted]). Kitcher's approach also faces a number of other difficulties (Barnes 1992; Woodward 2003, 358-373).

What are we to make of these problems? Many have taken them to be indicative of the demise of deductivist approaches to explanation. The D-N model was beset by problems, and Kitcher attempted to respond to these within a deductivist, non-causal framework. The intuitive idea that explanation and understanding have something to do with science's ability to unify our account of empirical phenomena is attractive. But deductivists have been unsuccessful at developing a systematic account of this intuition. The consensus seems to be that the failure to develop an adequate account of explanation in terms of unification reveals the sterility of deductivist approaches. I believe this is an entirely mistaken conclusion to draw. What the problems with Kitcher's account show is not that a deductivist approach articulating global conditions on explanation is a bad strategy, but only that the program of explanatory unification is unsuccessful in doing so.

A complete account of scientific explanation must provide an analysis of this fundamental feature of science, namely that certain global theories have independent explanatory power. With respect to many well-confirmed theories of broad scope, it is widely accepted that they have independent explanatory power. Of course, the consensus may be mistaken about any particular global theory: that it is well-confirmed, that its scope is broad, or that it has independent explanatory power. Indeed, with respect to some global theories there may be no consensus at all among scientists or philosophers (as there are divergent views about what, if anything, quantum theories explain). Whether or not some global theory has independent explanatory power—and can play the role of underwriting the explanatory value of laws or dependency relations figuring



in an explanans—is a key question for any account of explanation in science, but it is not one that can be answered entirely in syntactic or formal terms, as Kitcher attempts to do. Causal approaches to explanation are generally committed to some version of the idea that certain global theories are explanatory because they correctly represent causal structure. Deductivists have sometimes taken a neo-Kantian approach, on which causal relations and explanatory power emerge in the limit of the rational progress of science towards maximally unified theories (Friedman 1974; Kitcher 1989, 497-499). At this point, we must leave to one side these difficult and interesting issues. The remainder of the present paper focuses rather on the prior global condition on explanation, namely the nature of the connection between explanatory global theories and local models.

#### **4. Causal, minimal model and structural accounts**

Much recent work on scientific explanation focuses on the role of idealized explanatory models (Woodward 2003; Strevens 2008; Batterman 2010; Bokulich 2011; Batterman and Rice 2014; Healey 2015; Reutlinger 2016; Saatsi 2016). This work has been successful in enriching our understanding of explanation in areas of science where earlier accounts—whether causal or deductivist—are inapplicable. I believe this work has largely overlooked or misunderstood the role of global theory in explanation, and the connection between global theory and local model, and it has thereby misconstrued how idealized models underwrite scientific explanation. This section discusses three prominent accounts of model-based explanation: Woodward’s causal difference-making account, Batterman and Rice’s minimal models account, and Bokulich’s structural account (Woodward 2003; Batterman and Rice 2014; Bokulich 2011). Woodward, Batterman and Rice include no role for global theory in their accounts, and they deny that anything like theoretical integration is necessary for explanation. Bokulich does discuss the justification of models “top-down” from theory, but she does not flesh out how this justification works to underwrite explanation.

For Woodward, what makes a causal generalization explanatory is that it can answer a relevant range of “what-if-things-had-been-different” questions, and it does this by supporting difference-making counterfactuals about what would happen under scientifically relevant interventions on the explanans variables. For example, consider the derivation, in the context of an electrostatic model, of the electric force on a charged object near an infinitely long, uniformly charged wire; this is one of the main examples Woodward uses to introduce and motivate his account of causal explanation (Woodward 2003, 187-194). Woodward cites Coulomb’s law and describes the standard textbook explanation of the result. What makes this a causal explanation, according to Woodward, is that Coulomb’s law is an explanatory generalization that can answer the right sort of “what-if-things-had-been-different” questions.

Although the Coulomb derivations correctly represent causal relations according to Woodward’s criteria, and these derivations satisfy Woodward’s conditions for causal

explanation, the derivations nevertheless are not explanatory, at least not without support from an explanatory global theory ([self-reference omitted]). In 1785, Coulomb performed measurements of electrostatic force using a torsion balance. He determined the inverse power of  $r$  that appears in Coulomb's law to be  $2 \pm 0.02$ . Coulomb and his contemporaries used this regularity to make predictions and answer what-if-things-had-been-different questions about a range of experimental setups in electrostatics, including the sorts of wires, coils and charged plates Woodward considers. Their predictions of experimental outcomes were correct, their answers to what-if-things-had-been-different questions in electrostatics were correct, and they had a correct representation of at least part of the actual causal structure of the experimental setup. Yet my claim is that Coulomb and his contemporaries did not explain these results. Coulomb's account is what physicists would call a phenomenological model, one based solely on experimental regularities and applicable to a narrow range of similar physical systems. These models are useful for predictions, they map dependency relations and sometimes map local causal structure, and they often have great heuristic power—but they are not explanatory.

As I have argued elsewhere, explanations of electrostatic phenomena begin with Maxwell, almost a hundred years later. Maxwell's electromagnetic theory is founded on fundamental physical principles, including charge conservation and energy conservation. The time-dependent electromagnetic field introduced by Maxwell stores energy and plays an active causal role; it is a basic constituent of many physical systems. Electromagnetic theory has broad empirical scope, providing a unified description of all classical electromagnetic and optical phenomena. The theory is very well confirmed. We have good reason to believe that Maxwell's equations accurately describe the classical electromagnetic field and electromagnetic phenomena. Indeed, Coulomb's law for electrostatics can be derived from and explained by electromagnetic theory. It is only within electromagnetic theory, for instance, that the concept of the energy of a system of electrostatic charges can be defined and determined. It is this connection to electromagnetic theory that renders the derivations containing Coulomb's law explanatory. Electromagnetic theory explains why Coulomb's law is true, and in fact why the inverse power of  $r$  must be exactly 2. Electromagnetic theory also provides the theoretical underpinning for explanations of electrostatic phenomena by, for example, showing that electrostatic systems contain not only point charges, but also an electromagnetic field with a well-defined energy density. The electrostatics case motivates my contention that Woodward neglects the essential feature of explanatory integration with global theory, an argument I develop more fully in ([self-reference omitted]).

Batterman and Rice's account of explanation faces a similar challenge. They focus on one kind of highly idealized model, minimal models, in which a macro phenomenon is exhibited across a range of diverse micro realizers. Minimal models are explanatory, they argue, because of a backstory about why the features that distinguish detailed models exhibiting the explanandum phenomenon—features omitted from the minimal model—are irrelevant to the

occurrence of that phenomenon. According to Batterman and Rice, it is one thing to explain *that* a given detailed model exhibits a certain behaviour. Minimal model explanations go farther, telling us why and under what conditions various detailed models exhibit this behaviour and thereby “provide the deeper understanding we seek” (Batterman and Rice 2014, 375). Batterman and Rice are absolutely right to stress this important aspect of explanation via idealized models, one that has largely been overlooked in the literature. My contention is that the backstory given by Batterman and Rice is incomplete. They claim to provide “an analysis of what actually does the explanatory work” in explanations invoking minimal models (Batterman and Rice 2014, 349). Their analysis, however, ignores one key component doing a great deal of explanatory work: the theoretical integration of the minimal model with global theory.

Batterman and Rice consider R. A. Fisher’s model of the sex ratio in a biological population. This is a minimal model in which the details that distinguish diverse biological populations are omitted and the relevant dynamics are all encoded into a single variable, the “substitution cost” for producing male and female offspring. Fisher’s explanation of the 1:1 sex ratio in a population appeals to the fact that this substitution cost is linear in the population, which means that males and females cost the same amount of resources to produce. Batterman and Rice aim to give an account of how Fisher’s minimal model can answer the key question, “Why do very different populations have these features [1:1 sex ratio] in common?” (Batterman and Rice 2014, 370). It is not enough simply to say that very different populations have the same substitution costs; as Batterman and Rice emphasize, “[w]ere we simply to cite the fact that all these populations have the common feature of linear substitution cost, we would fail to explain this universal behavior” (Batterman and Rice 2014, 373). Rather, “we need a story about why a linear substitution cost will be found in such diverse systems” (Batterman and Rice 2014, 370).

According to Batterman and Rice, such a story consists of demarcating a group of detailed models exhibiting the 1:1 sex ratio, called the “universality class” of such models. This involves showing that the models exhibit equilibrium behaviour: one can change certain variables in these models without perturbing the sex ratio (e.g., variables that distinguish a population of sheep from a population of mule deer). Biologists have been successful in demonstrating the robustness of this equilibrium behaviour across detailed models of a range of biological systems and a range of genetic assumptions. Fisher’s minimal model exhibits the same robustness. For Batterman and Rice, Fisher’s model is not explanatory by virtue of accurately representing biological systems, by sharing common features with these systems, or by tracing the causal structure of these systems (since the model does none of these things). Rather, Fisher’s model is explanatory because biologists have derived the result—for a group of models which *are* detailed and *do* accurately represent target biological systems—that the dynamical processes represented in these detailed models leads to linear substitution cost. Further, for the same group of models, the linear substitution cost does not depend on, and is invariant under changes to, a set of variables representing micro details. Fisher’s model has the feature of linear substitution cost, and it is also

invariant under changes to these micro details (indeed, it lacks variables representing these micro details). There is thus a universality class which includes the group of detailed models of target systems of interest, and Fisher's minimal model, all sharing these characteristics. Fisher's model is explanatory because it is a member of this class.

My contention is that this backstory must include something that Batterman and Rice omit: an account of the theoretical integration of the model with an explanatory global theory. The story they give about why Fisher's model is explanatory, briefly summarized in the previous paragraph, seems to me correct as far as it goes. However, their account of minimal model explanation does not provide a *criterion* for distinguishing satisfactory stories—ones demonstrating the explanatory power of the local model—from unsatisfactory ones. Their emphasis on identifying an explanatory idealized model as a member of a universality class does not constitute a satisfactory account of what makes the idealized model explanatory. It is not enough that for Fisher, “identifying the commonality of the linear substitution cost allows us... to explain the 1:1 ratio across many natural populations” (Batterman and Rice 2014, 367). Rather, what renders Fisher's minimal model explanatory is the additional fact, not contained in the explanans, that linear substitution cost in the detailed models is *itself* theoretically justified by contemporary evolutionary biology. Biologists give a fitness-based justification of why so many populations exhibit linear substitution cost. Part of this account is that under normal conditions, when a population moves away from a 1:1 sex ratio, there is a fitness advantage for parents who produce the minority sex and thus a non-linear substitution cost. This leads to overproduction of the minority sex and pushes the ratio back towards equilibrium, where the fitness advantage disappears and the substitution cost is linear.

To see this another way, consider a version of creation science that accepts many of the results of contemporary evolutionary biology, including Fisher's model, but rejects speciation and large-scale evolution by natural selection. Our hypothetical creation scientists work with the same universality class of detailed population models as do biologists, with the same features of linear substitution cost and stable equilibrium around a 1:1 sex ratio. There is the same mathematical story about why the various micro features shared by detailed models in the group are irrelevant to their macro behavior, and about why Fisher's model is a member of the same universality class as those models in the group. Creationists can show that the 1:1 ratio is stable under various changes to gene frequencies, mating strategies, sexual vs. asexual reproduction, etc. They can equally well prove, rather than assume, stability of macro behaviour under extensive changes to micro variables. So according to the account of minimal model explanation given by Batterman and Rice, Fisher's model would count as explanatory in this context.

But it is not. The reason Fisher's model is not explanatory in creation science is that it is not appropriately integrated with a global scientific theory having independent explanatory power. More specifically, while the feature of linear substitution cost is present in the relevant detailed models of target systems, the feature is not theoretically justified, that is, justified by a global

scientific theory. Biologists and creationists appeal to the same minimal model and put forward the same argument as putative explanations. The local features of the putative explanation, such as the truth of the premises and its inferential structure, are the same. However, in the biologists' case this argument is explanatory and in the creation scientists' case it is not (or at least, not a scientific explanation). The difference lies in the global features, and in particular the theoretical integration of features of the idealized model appealed to in the explanation with a global theory that has independent explanatory power. Creationists do have a global story: so many populations exhibit linear substitution cost because God wants it to be so, and God created ancestor populations with linear substitution costs. But it is not a global theory, because it is not a scientific theory. As I shall argue, theoretical integration with a global theory is a necessary condition for scientific explanation via idealized models (Section 5).

But why not count Fisher's model as explanatory in the context of creation science, and Coulomb's model as explanatory prior to Maxwell? These local models perhaps explain a narrower range of phenomena than models supported by global theory, the objection goes, but they still qualify as explanatory. An explanans describing a local model can genuinely explain an explanandum regardless of whether or not we have the correct explanation of why the explanans itself obtains. This is certainly the case in our non-scientific, common-sense explanations.

Keep in mind that our focus here is understanding how scientific explanation functions using highly idealized models in science. We should expect that the requirements for successful explanation in this context to differ significantly from those of everyday or religious explanation. One issue is that explanations appealing to idealized scientific models require particularly high standards—a point not adequately acknowledged in the model-explanation literature. These are explanations where we know some of the explanans is false, because it describes objects, properties or dependency relations that do not exist in the physical system. Of course, not just any false statements that entail an explanandum qualify as explaining it. Scientists work with idealized models to make predictions about phenomena of interest. In some cases, they claim the models do not explain the predicted phenomena (often calling these phenomenological or data models). In other cases, they claim the models do provide scientific explanations. The focus on global theory, and in particular the criterion of theoretical integration developed in Section 5, makes sense of many of these scientific judgments. Accounts of scientific explanation aim for a reflective equilibrium between philosophical account and the best judgments of the relevant scientific community. There is a circularity in this approach, to be sure, but it is a circularity present in all successful philosophical accounts of scientific method.

In contrast with the rest of the literature on model-based explanation, Bokulich's approach foregrounds the role of global theory in justifying explanatory uses of local models. Our focus here is on this component of her account, what she calls the "justificatory step" of showing that the model is a good one within its domain of applicability.

Very broadly, we can understand this justificatory step as specifying what the domain of applicability of the model is, and showing that the phenomenon in the real world to be explained falls within that domain. This justificatory step is intended to call explicit attention to the detailed empirical or theoretical process of demonstrating the domain of applicability of the model (Bokulich 2011, 39; for a comprehensive assessment of Bokulich's structuralist account of model-based explanation, see King 2015).

A model is justified, according to Bokulich, when it can satisfactorily capture explanatorily relevant features of its target system. This is different from mere empirical adequacy, since the goal of this criterion is to rule out an explanatory role for data or phenomenological models that are predictively accurate. Rather, justification requires demonstrating "where and to what extent the model can be trusted as an adequate representation of the world" (Bokulich 2011, 39). This justification may proceed top-down, where a global theory supports the local model. "More frequently, however, there is no overarching theory," and in the absence of global theory justification proceeds bottom-up (Bokulich 2011, 39). Here, the idealized elements of the model can be reliably de-idealized so that features of the model are shown to be accurate representations of corresponding features of the target.

Bokulich's work has been helpful in broadening our focus on the role of global theory in model-based explanation, but I think her characterization of the "justificatory step" is flawed, and for two reasons. First, it is not the case that justification can proceed exclusively "bottom-up," as she claims. Bokulich provides no argument that bottom-up justification can be sufficient in the absence of global theory, and in fact she speculates that there may be a top-down component in all cases (Bokulich 2011, 39, fn. 12). Most tellingly, she does not give any examples of exclusively bottom-up justification. Her key example of an explanatory idealized model, Bohr's model of the atom, includes top-down justification. Bohr's model was developed in 1913, prior to any global theory of the quantum domain. Bokulich claims the model is justified top-down by the subsequent development of contemporary semi-classical mechanics (Bokulich 2011, 43; Bokulich 2008, Chapter 5).

Second, Bokulich's description of the "justificatory step" is simply too vague to do the work she claims it can do of distinguishing explanatory models from merely phenomenological ones. As we have seen, in the explanatory practice of contemporary science the connections between global theory and local model are diverse and complex. Some connections underwrite the explanatory power of the local model, and others do not. What is needed is a detailed account of the necessary conditions for these connections to support explanation—the relation I call theoretical integration—and this is something Bokulich does not provide. For instance, she claims that semi-classical mechanics provides a top-down justification of the explanation of the spectrum of hydrogen by Bohr's model. An adequate account of model-based explanation needs to provide the resources, in the form of a clear criterion, to substantiate such a claim. Unfortunately, Bokulich (along with many others) seems to confuse this issue with the distinct

question of explanatory depth. It is one thing to embrace explanatory pluralism, as Bokulich does, and claim that some explanations of a given phenomenon may be deeper or better than others; so the explanation offered by Bohr's model is not as deep as the one offered by quantum mechanics (Bokulich 2011, 44). It is another matter entirely to claim that the derivation of the spectrum of hydrogen from Bohr's model is explanatory at all, in the sense that it satisfies well-defined conditions on explanation via local models. The latter claim requires distinguishing explanatory models from phenomenological ones, and this in turn necessitates having a clear account of the conditions under which a local model is theoretically integrated with global theory. To this we now turn.

## 5. Theoretical integration

The account of explanatory integration put forward here contains two global necessary conditions on explanation: that there be a connection between global theory and local model such that the global theory confers explanatory power on the local model, and that the global theory itself have independent explanatory power. The remainder of this paper focuses on the former condition, which I dub *theoretical integration* (TI). This section provides two examples of TI, first in the explanation of the period of a simple pendulum, and second in the explanation of gravitational waves from distant stellar objects. It then sketches an account of TI, and applies it to the models of Fisher and Bohr considered earlier.

Ernan McMullin first proposed that certain idealized models, which he calls Galilean, can underwrite scientific explanation (McMullin 1985). Galilean idealization is characterized by the fact that the idealized model approximates the target system and, more importantly, that complementary to idealization are reverse techniques for adding back real-world details and de-idealizing by eliminating simplifying assumptions. Galilean idealizations thus have an intrinsic "self-correcting" feature such that they can (at least in principle) be brought in ever closer agreement with empirical observations in a theoretically justified, non-*ad hoc* way. Others have called these sorts of idealizations "controllable" or "harmless" (Elgin and Sober 2002; Batterman 2005, 235; [self-reference omitted]).

McMullin's paradigm case of explanation via Galilean idealization is Galileo's pendulum. My claim is that a necessary condition for the derivation of the period of a pendulum using Galileo's pendulum law to be explanatory is that the idealized pendulum model, of which Galileo's pendulum law is true, be theoretically integrated with a global theory having independent explanatory power, such as Newtonian mechanics. In this case, the local pendulum model is theoretically integrated with global theory directly, without intermediary or additional support from related models. In a somewhat metaphorical sense, the local model is approximately a submodel of the global theory. Call this *direct TI*: the top-down justification of a Galilean local model by a global theory. It is important to keep in mind that the issue of theoretical integration is independent of the issue of the explanatory power of the global theory.

My claim here is only that Galileo's pendulum model exhibits direct TI; whether the model supports bona fide scientific explanations depends on whether Newtonian Mechanics has explanatory power and on a host of local conditions. Thus, one may not agree with McMullin that Galileo's pendulum law is explanatory, or that Newtonian Mechanics has independent explanatory power, yet agree that the local model is still directly integrated with global theory.

As we have seen, satisfactory account of explanation in science needs to accommodate the fact that many explanatory models are not Galilean and are not directly integrated with global theory. Non-Galilean models contain idealizations that are not self-correcting, and the idealizations cannot be said to correspond to elements of the target system. The model contains elements that support good scientific explanations, yet these elements cannot plausibly be regarded as representing features of the target. Non-Galilean models contain idealizations that are not corrigible, in the sense that it is not possible systematically to refine the idealization to bring these elements in closer and closer agreement with the target system. These models may be explanatory by virtue of *indirect TI*. Here, the model is explanatory because it is appropriately related to one or more other local models that are Galilean, where the latter models are directly integrated with global theory. As well, it may be indirectly justified by relation to the global theory. This more complicated situation involves sideways justification from one or more related local models (themselves integrated with global theory) and may also involve some degree of top-down justification from a global theory.

The central role played by non-Galilean models in the explanatory activity of contemporary science, the relation of theoretical integration, and the indirect contributions of related models may be illuminated with an example. The case is the explanation of gravitational waves from distant stellar objects using the general theory of relativity. What makes the case interesting is that there are two distinct routes to predicting the same gravitational wave phenomena in general relativity, one of which is explanatory and the other of which is not. Further, the difference in the explanatory value of the two approaches is due to the fact that one approach is theoretically integrated with the global theory of general relativity and the other is not (see [self-reference omitted] for an extended analysis of this case).

Soon after formulating general relativity, Albert Einstein predicted the possibility of gravitational waves: short-wavelength perturbations of the spacetime metric approximating plane waves. Over the last three decades, gravitational physics has developed accurate, quantitative models of gravitational waves emitted from realistic astronomical sources (Blanchet 2014), and gravitational waves were recently detected for the first time (Collaboration, Virgo et al. 2016). One of the main objects of study are inspiral systems composed of two neutron stars or black holes orbiting each other, spiraling inward and emitting gravitational radiation. There are two routes to deriving experimentally-measurable gravitational wave phenomena, with comparable predictive power but very different explanatory value. The traditional approach is to construct a post-Newtonian (PN) model of the inspiral system by starting with the fundamental equations of



general relativity and making a number of idealizing assumptions. One such idealization is that the local PN model contains point-particle sources with finite mass and infinitesimal volume. Derivations of gravitational wave phenomena appealing to the PN model contain lawlike premises describing dependency relations in the model that are inconsistent with the fundamental equations of general relativity. More recent effective field theory (EFT) models of the same inspiral system offer even more efficient derivations of gravitational wave phenomena. Here again, we start with the fundamental equations of general relativity and make a series of idealizing assumptions to arrive at the local EFT model.

The predictions of the two approaches match, but the assumptions and methods—and the explanatory merits—are very different in PN and EFT cases. The difference in the explanatory merits of PN and EFT models has to do with the ways in which these local models are constructed. Both PN and EFT models of inspiral systems make use of a point-particle idealization. But point-particle sources are highly problematic within the global theory of general relativity. Point particles are physically impossible according to general relativity, and they are mathematically untreatable by the fundamental equations of the theory. This makes both PN and EFT models non-Galilean. They both contain an element that does not approximate the target system, one that cannot be reverse-engineered, and one that is not physically justified—indeed, it is physically impossible according to global theory.

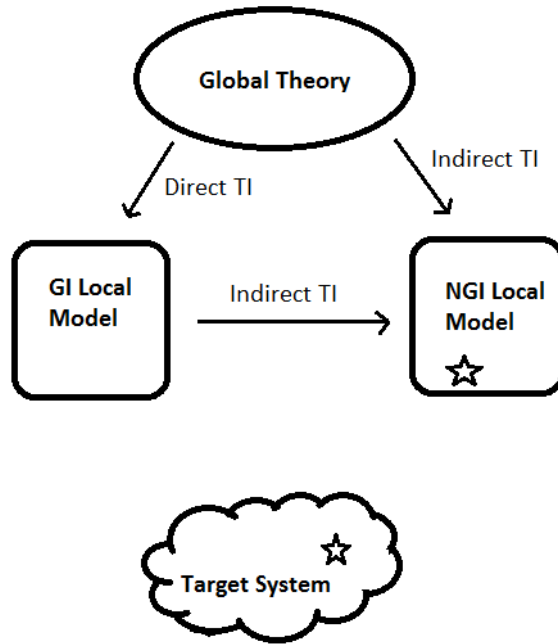
The key point for our purposes is that there are significant differences in the role of this idealization in the two models. In the PN model, the point-particle idealization can be given some justification, briefly as follows. It is possible to construct other local models of inspiral systems that are similar to the PN models yet treat the objects as extended bodies instead of point particles. These related models yield a greatly restricted range of predictions compared with the point-particle PN model. A reasonably good argument, one persuasive to reflective physicists in the field, can be made that the point-particle idealization will be able to be discharged entirely in some future PN models with empirical scope and accuracy comparable to current point-particle PN models. PN models are constructed based on solutions of the fundamental equations of motion of general relativity. While these are not regular perturbative solutions, and they involve contrary-to-fact assumptions, the claim is that there is a reasonably strong theoretical justification of the model-construction methods.

Accurate higher-order PN predictions of gravitational waves have been available since the mid-1980s. It is only since the development of extended-body PN models in the mid-2000s, however, that the non-Galilean point-particle idealization has been shown to be potentially dischargeable. Today these predictions also constitute *explanations* of observed gravitational phenomena. In other words, the PN model is theoretically integrated with the global theory of general relativity because the justification of particular non-Galilean idealization used to construct the model is reasonably strong (this claim is developed and defended in [self-reference omitted]). By contrast, the point-particle idealization in EFT models is not justifiable. The

idealization is a required assumption: point particles are the basic building blocks of effective field theories. For this reason, the point-particle assumption cannot be discharged in the EFT model. There is no theoretical integration with general relativity because the point-particle assumption is essential and ineliminable in EFT models. Thus, derivations of gravitational waves appealing to these models are predictive but not explanatory. While the predictions of the two approaches closely match, the assumptions and methods—and the explanatory merits—are very different in PN and EFT models.

This example of the explanation of gravitational waves via PN models introduces one key feature of theoretical integration, the justification of idealizing steps in model construction. Here, the justification comes from two directions: top-down from general relativity and sideways, so to speak, from related extended-body models. In this case, theoretical integration was achieved by connecting PN models to related local models that were free from a particularly problematic idealized element.

Figure 1 depicts an example of indirect TI. A global theory is depicted at the top of the diagram. At the bottom is the target system containing the phenomenon to be explained, illustrated with a star. In the middle are two local models of the target system, one Galilean (GI) and the other non-Galilean (NGI). The NGI model contains an element, indicated with a star, that accurately represents the explanandum phenomenon. The GI local model represents the target as well, but does not contain an accurate representation of the explanandum phenomenon in the target. A necessary condition for the NGI model to be explanatory is that it be theoretically integrated with a global theory, and in this case this integration flows indirectly from the GI model and the global theory. Figure 1 illustrates a simplified version of the PN model case, where the global theory is general relativity, on the left is an extended-body PN model, and on the right is a point-particle PN model.



*Figure 1.* Theoretical integration between a global theory, a Galilean local model, and a non-Galilean local model.

It should be emphasized that Figure 1 illustrates one configuration of indirect TI; in other cases a local model may be theoretically integrated along very different sorts of pathways. For instance, Fisher’s minimal model is non-Galilean and satisfies indirect TI. The model is explanatory of a regularity, namely the universality of a 1:1 sex ratio across a range of target systems with different micro realizers. Theoretical integration comes from connections the model has with detailed models that are Galilean and are theoretically integrated with contemporary evolutionary theory. Here, indirect TI flows exclusively from related Galilean local models (the collection of detailed models in the same universality class), and not from global theory.

This account of global conditions on explanation is far from complete. However, it already provides us with a framework for understanding the pathways of justification from global theory to explanatory local model, an improvement on prior accounts of model-based explanation. As we have seen, Bokulich’s account includes a “justificatory step” demonstrating that the model is “an adequate representation of the world within its domain of applicability,” and “a good model, able to adequately capture the relevant features of the world” (Bokulich 2011, 39). In order to distinguish explanatory from non-explanatory local models, on her account, we need to determine whether a model is an adequate representation of the target system. One might worry about tying explanatory power to representational adequacy in the context of idealized models; in any case, the criterion is too vague to do the work needed. My account fills in several of the key pieces: a local model is explanatory when it bears particular derivational relations to a global theory, and in some cases to a related Galilean local model.

My account also sheds light on Bokulich's claim that Bohr's model explains the spectrum of hydrogen. If I am right, a necessary condition for Bohr's model to underwrite explanations is that it be theoretically integrated with an explanatory global theory. As we have seen, Bokulich considers semi-classical mechanics to be the relevant global theory, and she shows that Bohr's model is justified top-down by its connection with contemporary semi-classical mechanics (Bokulich 2008, Chapter 5). To my mind, it is implausible that semi-classical mechanics qualifies as a global theory with independent explanatory power. At the very least, much more needs to be said in support of the claim that it is. Assuming that semi-classical mechanics does not have independent explanatory power, let's consider quantum mechanics as the relevant global theory (setting aside the vexed question of whether quantum mechanics explains!). The question remains whether Bohr's model is theoretically integrated with quantum mechanics. On the one hand, Bohr's model does not satisfy direct TI. Definite electron trajectories, which are essential to the Bohr model, are impossible according to quantum mechanics. This is very much like the EFT case in general relativity, where an ineliminable idealization in the local model blocks explanation. On the other hand, there are no related local models to provide indirect TI. Quantum mechanical models of the hydrogen atom cannot be related to, or integrated with, the Bohr model. More sophisticated local semi-classical models of the hydrogen atom are directly related to the Bohr model, but again these are not integrated with the global theory of quantum mechanics. In sum, the condition of theoretical integration developed here provides good reasons to be sceptical that Bohr's model is an explanatory model. More generally, sharpening our focus on global conditions gives us the tools to identify cases, widespread among philosophers and scientists alike, of non-explanatory idealized models for which explanatory virtues are claimed.

## 6. Conclusion

The account of explanatory integration introduced here shows how scientific explanation functions in the context of idealized models by articulating certain local and global necessary conditions that the explanation and the models must satisfy. The local conditions are that statements in an explanation be true of a model, and that the explanans include dependency relations. The global conditions are that there be a global theory with independent explanatory power, and the local model be theoretically integrated with the global theory.

The main focus has been on the neglected role of global theory in scientific explanation and the key condition of theoretical integration between model and theory. Proposed explanations of particular regularities or phenomena, such as a 1:1 sex ratio or gravitational waves, sometimes make essential use of highly idealized, non-Galilean local models. These derivations contain law statements describing dependency relations in a local model, statements that are in some cases inconsistent with the global theory. A necessary condition for these derivations to be explanatory is that the local model be theoretically integrated with the global theory, for only then can the local model support explanation. I have shown how this account can overcome some of the

problems facing other recent accounts of model-based explanation. As we have already seen, the configuration of local models, the relations of direct and indirect TI between them, and their relations with global theory are diverse and complex. Clearly, further analysis and detailed case studies are needed to develop a more robust account of the basic concept of TI introduced here.

The virtues of explanatory integration are precisely what Hempel and Oppenheim took to be essential to a broadly empiricist approach to explanation in science. Local conditions of adequacy stipulate characteristics of the explanans and the relation between the explanans and the explanandum, most importantly that the explanandum be the conclusion of a special kind of deductive argument. Global conditions underwrite the explanatory role of laws or dependency relations in the explanans by connecting them with an appropriate global theory. The account aims to provide a comprehensive treatment of causal explanation, non-causal explanation and—not considered in 1948—explanations making use of highly idealized models.

## Acknowledgements

[To be provided]

## References

- Batterman, Robert and Collin Rice. 2014. "Minimal Model Explanations." *Philosophy of Science* **81**(3): 349-376.
- Batterman, Robert W. 2002. *The Devil in the Details: Asymptotic Reasoning in Explanation, Reduction, and Emergence*. Oxford: Oxford University Press.
- . 2005. "Critical Phenomena and Breaking Drops: Infinite Idealizations in Physics." *Studies in History and Philosophy of Modern Physics* **36B**(2): 225-244.
- . 2010. "On the Explanatory Role of Mathematics in Empirical Science." *British Journal for the Philosophy of Science* **6**(1): 1-25.
- Blanchet, Luc (2014) "Gravitational Radiation from Post-Newtonian Sources and Inspiralling Compact Binaries." *Living Reviews of Relativity* **17** DOI: 10.12942/lrr-2014-2.
- Bokulich, A. 2011. "How Scientific Models Can Explain." *Synthese* **180**(1): 33-45.

- Bokulich, Alisa. 2008. *Reexamining the Relationship between Classical and Quantum Mechanics: Beyond Reductionism and Pluralism*. Cambridge: Cambridge University Press.
- . 2009. "Explanatory Fictions." *Fictions in Science : Philosophical Essays on Modeling and Idealization; Routledge Studies in the Philosophy of Science* ;. Mauricio Suárez. New York, NY: Routledge.
- Collaboration, Ligo Scientific, Collaboration Virgo, et al. 2016. "Observation of Gravitational Waves from a Binary Black Hole Merger." *Physical Review Letters* **116**(6): 061102.
- Elgin, M. and E. Sober. 2002. "Cartwright on Explanation and Idealization." *Erkenntnis* **57**(3): 441-450.
- Friedman, Micheal. 1974. "Explanation and Scientific Understanding." *Journal of Philosophy* **71**(1): 5-19.
- Giere, Ronald. 1988. *Explaining Science: A Cognitive Approach*. Chicago, IL: University of Chicago Press.
- . 2009. "Why Scientific Models Should Not Be Regarded as Works of Fiction." *Fictions in Science : Philosophical Essays on Modeling and Idealization; Routledge Studies in the Philosophy of Science* ;. Mauricio Suárez. New York, NY: Routledge.
- Healey, Richard. 2015. "How Quantum Theory Helps Us Explain." *The British Journal for the Philosophy of Science* **66**(1): 1-43.
- Hempel, Carl G. 1965. "Aspects of Scientific Explanation." *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*. New York: Free Press: 331-496.

- Hempel, Carl G. and Paul Oppenheim. 1948. "Studies in the Logic of Explanation." *Aspects of Scientific Explanation, and Other Essays in the Philosophy of Science*. New York: The Free Press: 245-290.
- Kim, Jaegwon. 1963. "On the Logical Conditions of Deductive Explanation." *Philosophy of Science* **30**: 286-291.
- King, Martin. 2015. "On Structural Accounts of Model-Explanations." *Synthese*: 1-18.
- Kitcher, Philip. 1981. "Explanatory Unification." *Philosophy of Science* **48**: 507-531.
- . 1989. "Explanatory Unification and the Causal Structure of the World." *Minnesota Studies in the Philosophy of Science, volume XIII*. Philip Kitcher and Wesley C. Salmon. Minneapolis: University of Minnesota Press: 410-506.
- McMullin, Ernan. 1985. "Galilean Idealization." *Studies in History and Philosophy of Science* **16**(3): 247-273.
- Reutlinger, Alexander. 2016. "Is There a Monist Theory of Causal and Non-Causal Explanations? The Counterfactual Theory of Scientific Explanation." *Philosophy of Science* **Forthcoming**.
- Saatsi, Juha. 2016. "On Explanations from Geometry of Motion." *The British Journal for the Philosophy of Science*.
- Strevens, Michael. 2008. *Depth : An Account of Scientific Explanation*. Cambridge, Mass.: Harvard University Press.
- Woodward, James. 2003. *Making Things Happen : A Theory of Causal Explanation*. Oxford: Oxford University Press.